

## 大数据视角分析学习变革\*

## ——美国《通过教育数据挖掘和学习分析促进教与学》报告解读及启示

徐 鹏<sup>1</sup> 王以宁<sup>2</sup> 刘艳华<sup>3</sup> 张 海<sup>2</sup>

(1.东北师范大学 信息化管理与规划办公室;

2.东北师范大学 传媒科学学院;

3.吉林大学 公共外语教育学院,吉林长春 130012)

[摘 要] “数据驱动学校,分析变革教育”的大数据时代已经来临,利用教育数据挖掘技术和学习分析技术,构建教育领域相关模型,探索教育变量之间的相关关系,为教育教学决策提供有效支持将成为未来教育的发展趋势。“大数据”的出现,将掀起人类教与学的又一次变革。2012年,美国国家教育部发布了《通过教育数据挖掘和学习分析促进教与学》报告,对美国国内大数据教育应用领域和案例,以及应用实施所面临的挑战进行了详细的介绍。借鉴此报告,我们认为未来我国教育领域的大数据研究和应用,应加强国家和地方对相关的研究和应用和技术层面、管理体制层面以及法律制度层面的支持,按需合理规划具体研究和应用,整合现有资源,发挥后发优势,借助“大数据”实现真正意义上的个性化学习,进而实现教育公平。

[关键词] 大数据;教育数据挖掘;学习分析

[中图分类号] G434

[文献标识码] A

[文章编号] 1672—0008(2013)06—0011—07

## 一、报告发布背景简介

近年来,随着互联网的飞速发展,人们越来越多的行为在网络中发生,这直接导致互联网中人类行为相关数据呈爆炸式增长,人类在不知不觉中已经进入了一个“大数据”时代。联合国在2012年发布的大数据白皮书“*Big Data for Development: Challenges & Opportunities*”中指出,大数据时代已经到来,大数据的出现将会对社会各个领域产生深刻影响<sup>[1]</sup>。为此,世界各国都在此领域投入大量人力物力来支持“大数据”相关研究和应用。美国作为世界科技强国,在2012年3月29日正式启动“Big Data Research and Development Initiative”计划,提出“通过收集、处理庞大而复杂的数据信息,从中获得知识和洞见,提升能力,加快科学、工程领域的创新步伐,强化美国国土安全,转变教育和学习模式”,正式将“大数据”提高到国家战略层面,美国政府为该计划投入了2亿美元<sup>[2]</sup>。2013年2月6日,美国国家科学基金会(National Science Foundation)宣布将额外投入1千万美元,以激励社会及

人文科学中的“大数据”研究的发展。另外,国际著名学术期刊 *Nature* 和 *Science* 分别在2008年和2011年推出大数据相关专刊,由此可见,“大数据”已经得到了世界各国学术界、工业界和政府机构的高度重视。

在教育领域,耶鲁大学、哈佛大学、斯坦福大学等世界知名高校也启动了教育大数据相关研究计划;另外,美国学校管理者协会(AASA)携手学校网络联合会(COSN),以及全球性的信息技术研究和咨询公司 Gartner 共同实施了一个名为“Closing the Gap: Turning Data into Action”的项目,旨在促进学校对学生信息系统和学习管理系统中大数据的使用。为了更好地促进美国国内“大数据”教育应用,为美国高等院校及 K-12 学校在“大数据”教育应用方面提供有效指导,美国教育部(U.S. Department of Education)在2012年10月发布了《通过教育数据挖掘和学习分析促进教与学》(*Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics*)报告(以下简称《报告》)。《报告》内容主要包括以下五个方

\* 基金项目:本研究系2011年度教育部高等学校博士学科点专项科研基金课题“信息技术深层整合教学结构与教师教育技术学科化研究”(课题编号:20110043110013)、东北师范大学哲学社会科学校内青年基金团队项目“文化公平观与完善公共文化服务体系建设研究”(课题号:12QN043)的系列研究成果之一,并受“中央高校基本科研业务费专项资金”资助。



面:个性化学习解读;教育数据挖掘和学习分析解读;自适应学习系统中大数据应用介绍;美国教育数据挖掘和学习分析应用案例介绍;美国的大数据教育应用挑战和实施建议。《报告》内容主要来源于以下三个方面:对公开发表或未公开发表的教育数据挖掘和学习分析相关文献的综述;对教育软件和学习管理系统开发公司中 15 位数据挖掘和分析方面专家的采访;8 位数据挖掘和学习分析领域学术专家讨论的总结<sup>[9]</sup>。

## 二、核心概念界定

### (一)教育大数据

“大数据”这一概念对于我们来说已不陌生,维基百科上给出的大数据定义是:数据量规模巨大到无法通过目前主流软件工具获取、管理、处理、整理成为帮助企业经营决策,达成更积极目的的信息<sup>[10]</sup>。

国际数据公司(IDC)认为,大数据是符合 4V 特征的数据集,即海量的数据规模(Volume)、快速的数据流转和动态的数据体系(Velocity)、多样的数据类型(Variety)、巨大的数据价值(Value)<sup>[11]</sup>。而《报告》则引用了国际著名的咨询公司麦肯锡(mckinsey)在 2011 年对大数据的定义,认为大数据是指数据量极大,以至于无法使用常规数据软件进行获取、存储、管理和分析的数据,“大数据”具有数据量大、数据多样和数据产生速度快三大特征<sup>[12]</sup>。教育领域中的大数据有广义和狭义之分,广义的教育大数据泛指所有来源于日常教育活动中人类的行为数据,它具有层级性、时序性和情境性的特征;而狭义的教育大数据是指学习者行为数据,它主要来源于学生管理系统、在线学习平台和课程管理平台等。

### (二)教育数据挖掘和学习分析

通过对教育大数据的获取、存储、管理和分析,我们可以构建学习者学习行为相关模型,分析学习者已有学习行为,并对学习者的未来学习趋势进行科学预测。《报告》指出,目前教育领域中大数据的应用主要有教育数据挖掘和学习分析两大方向,两个研究方向虽然同源,却在研究目的、研究对象和研究方法等方面截然不同,具体比较情况如表 1 所示。

表 1 教育数据挖掘和学习分析比较

概念名称 概念属性	教育数据挖掘	学习分析
研究目的	通过研究使学习者学习行为模型化显示,探寻各变量之间的相关关系,预测学习者未来学习发展趋势	通过研究使学校和教师能够根据不同学习者的能力水平和实际需求,为其提供合适的教育机会
研究对象	学习者学习行为相关数据	学习者学习行为、课程和学校运维等相关数据
研究方法	综合运用统计学、机器学习和数据挖掘的技术和方法	统计学、数据挖掘、社会学、心理学和学习科学的技术和方法

### 1.教育数据挖掘

教育数据挖掘是综合运用数学统计、机器学习和数据挖掘的技术和方法,对教育大数据进行处理和分析,通过数据建模,发现学习者学习结果与学习内容、学习资源和教学行为等变量的相关关系,来预测学习者未来的学习趋势。《报告》通过对教育数据挖掘领域专家进行访谈,列出了教育数据挖掘的四个研究目标<sup>[13]</sup>:(1)通过整合学习者知识、动机、元认知和态度等详细信息进行学习者模型的构建,预测学习者未来学习发展趋势。(2)探索和改进包含最佳教学内容和教学顺序的领域模型。(3)研究各种学习软件所提供的教学支持的有效性。(4)通过构建包含学习者模型、领域模型和教育软件教学策略的数据计算模型,促进学习者有效学习的发生。

《报告》指出,为了达到以上四个研究目标,研究者主要采用以下五类技术方法:(1)预测。建立一个能够从整合多个预测变量推断单一被预测变量的模式,例如,研究者通过在线学习环境中学习者参与在线讨论的情况、测试情况等,预测学习者在该门课程的学习中是否有失败的风险。(2)聚类。根据数据特性,将一个完整的数据集分成不同的子集,例如,研究者根据学习者在在线学习环境中学习困难、交互模式等将学习者分成不同的群组,进而为不同的群组提供合适的学习资源和组织合适的学习活动。(3)关系挖掘。探索数据集中各变量之间的相关关系,并将相关关系作为一条规则进行编码,例如,研究者利用关系挖掘,探索在线学习环境中学习者学习活动和学习成绩的相关关系,进而用于改进学习内容呈现方式和序列,以及在线教学方法。(4)人类判断过程简化。用一种便于人类理解的方式描述数据,以便人们能够快速判断和区分数据特征,该方法主要以可视化数据分析技术为主,用以改善机器学习模型。(5)模型构建。通过对数据集的聚类、相关关系挖掘等过程,构建供未来分析的有效现象解释模型。

### 2.学习分析

学习分析是近年来大数据在教育领域较为典型的应用,在国际上有专门针对学习分析研究和应用的国际会议“学习分析技术与知识国际会议”,目前已举办两届。在首届会议上将学习分析定义为“测量、收集、分析和报告有关学习者及其学习情景的数据集,以理解和优化学习及其发生情景”<sup>[14]</sup>。新媒体联盟(New Media Consortium)将学习分析定义为:利用松散耦合的数据收集工具和分析技术,研究分析学习者学习参与、学习表现和学习过程的相关数据,进而对课程、教学和评价进行实时修正<sup>[15]</sup>。我国学者顾小清认为,学习分析是围绕与学习者学习信息相关的数据,运用不同的分析方法和数据模型来解释这些数据,根据解释的结果来探究学习者的学习过程和情景,发现学习规律;或者根据数据阐释学习者的学习表现,为其提供相应的反馈从而促进更加有效的学习<sup>[16]</sup>。《报

告》中认为,学习分析是综合运用信息科学、社会学、计算机科学、心理学和学习科学的理论和方法,通过对广义教育大数据的处理和分析,利用已知模型和方法去解释影响学习者学习重大问题,评估学习者学习行为,并为学习者提供人为的适应性反馈。例如,教师和学校根据学习分析的结果,调整教学内容、对有学习失败风险的学生进行干预等。学习分析一般包括数据采集、数据存储、数据分析、数据表示和应用服务五个环节<sup>[1]</sup>。

### 三、报告核心内容解读

#### (一)自适应学习系统中教育大数据应用

在线学习系统中包含大量的细粒度的学习者学习行为相关数据,例如,学习日志、学习路径、学习成果数据、课程数据、学习管理数据等,研究者利用教育数据挖掘和学习分析技术对数据进行采集、存储和分析,为包括学校管理者、教师和学生在内的各级各类人员提供相应的反馈,以帮助其改善学校管理、教学和学习。《报告》为了让业内对大数据教育应用有一个系统完整的认识,以教育大数据在自适应学习系统中的应用为例进行了详细的说明<sup>[2]</sup>。

该自适应学习系统包含六大模块:(1)内容传递模块。管理、维护、传递个性化的学习内容与评价给学习者,以支持学习者的学习行为。(2)学习者数据库。存储学习者在学习系统中的时间戳标记的学习者输入和学习行为数据。(3)预测模块。整合系统外部学习者信息系统中的数据和系统内部学习者学习行为数据,通过对数据的处理和分析,对学习者的未来的学习行为和结果进行预测。(4)显示模块。将预测模块中的运行结果以可视化的方式显示给各类使用者。(5)自适应模块。根据预测模块的运行结果,触发内容传递模块,再根据学习者的学习水平和兴趣,推送合适的学习内容给学习者。(6)干预模块。允许教师、教学管理者和系统开发人员根据预测模块的运行结果,对自适应系统实施人为干预。

另外,自适应学习系统中除了这六大模块以外,还包括作为系统重要外部数据来源的学校、地区和机构的学生信息系统(SIS),该系统中存储着学习者相对稳定的个人基本信息数据。

基于大数据的自适应学习系统运行流程图如图1所示,第一步,学习者生成学习行为数据,经过内容传递模块,数据将被标记上时间戳;第二步,数据按照预先定义的结构存入学习者数据库;第三步,预测模块从学习者数据库和学生信息系统中采集数据,根据不同的分析目的,调用不同的分析工具和模型对数据进行分析;第四步,自适应模块根据预测模块中数据挖掘和分析的结果,通过内容传递模块为学习者提供合适的学习指导和学习策略;第五步,预测模块中数据挖掘和分析的结果同时被传递给显示模块,供教师和教学管

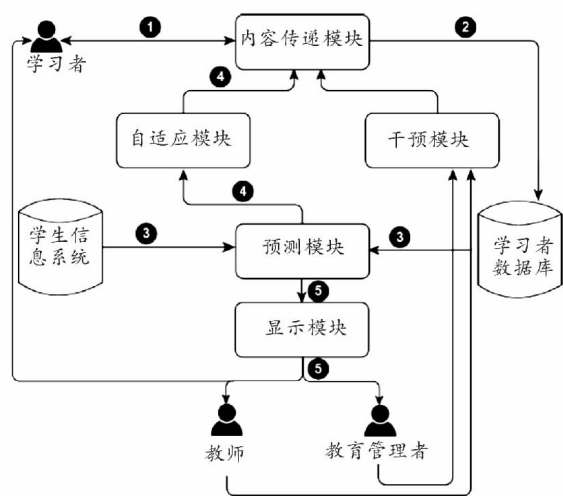


图1 基于大数据的自适应学习系统组成及运行流程

理者使用;最后,教师和教学管理者根据分析结果,通过干预模块对系统进行人为干预。

#### (二)教育数据挖掘和学习分析典型应用

《报告》通过对基于教育大数据的自适应学习系统的组成和运行流程的个案描述,使业内对于大数据的教育应用有了一个整体的认识,属于纵向描述;接下来,《报告》对教育数据挖掘和学习分析的应用领域作了详细描述,使业内对大数据的教育应用有了一个更加具体的认识,属于横向描述。《报告》中有关教育数据挖掘和学习分析应用领域的内容主要来源于对领域内公开发表和未公开发表文献的综述,以及对领域专家的访谈。

教育数据挖掘和学习分析应用领域主要包括:学习者的知识、行为和经历建模;学习者建档;领域知识建模;趋势分析。详细应用领域情况如表2所示。

表2 教育数据挖掘和学习分析应用领域<sup>[3]</sup>

应用领域	解决的问题	用于分析的数据
学习者知识建模	学习者掌握了哪些知识(例如:概念、技能、过程性知识和高级思维技能等)	①学习者正确的、不正确的和部分正确的应答数据;学习者作出应答花费的时间;帮助请求数据;犯错和错误重复数据 ②学习者的技能练习数据(内容和持续时间) ③学习者的测试(形成性和总结性)结果数据
学习者行为建模	学习者不同的学习行为范式与学习者的学习结果的相关关系	①学习者正确的、不正确的和部分正确的应答数据;学习者作出应答花费的时间;帮助请求数据;犯错和错误重复数据 ②学习者学习情境相关数据
学习者经历建模	学习者对于自己的学习经历的满意度	①满意度调查问卷和量表测试数据 ②在后续学习中学习者对于学习单元或课程的选择和表现数据
学习者建档	学习者聚类分组	学习者正确的、不正确的和部分正确的应答数据;学习者作出应答花费的时间;提示请求数据;犯错和错误重复数据





领域知识建模	学习内容的难度级别、呈现顺序与学习者学习结果的相关关系	①学习者正确的、不正确的和部分正确的应答数据;学习者在不同难度学习模块中的表现情况数据 ②领域知识分类数据 ③技能和问题解决的关联性数据
学习组件分析和教学策略分析	在线学习系统中学习组件的功能、在线教学策略与学习者学习结果的相关关系	①学习者正确的、不正确的和部分正确的应答数据;学习者在不同难度学习模块中的表现情况数据 ②领域知识分类数据 ③技能和问题解决的关联性数据
趋势分析	学习者当前学习行为和未来学习结果的相关关系	①在线学习系统中学习者学习行为相关的横向和纵向数据 ②学生信息管理系统中,持续一段时间且相对稳定的学习者基本信息数据
自适应学习系统和个性化学习	学习者个性化学习实现和在线学习系统自适应性实现	①在线学习系统中学习者学习行为相关的横向和纵向数据 ②与在线学习系统使用相关的用户反馈数据

### 1. 学习者知识建模

研究者通过采集学习者与在线学习系统的交互数据,包括学习者系统应答正确率、回答问题花费时间、请求帮助的数量和性质,以及错误应答的重复率等,这部分数据可以是课程层面的、学习单元层面的或知识点层面的。通过数据挖掘和分析,构建学习者知识模型,然后通过自动或人工反馈,为学习者在合适的时间,选择合适的方式,提供合适的学习内容。例如,Onsophic公司的在线学习平台就是通过收集平台中学习者每一个知识点学习(例如二次方程)的详细数据,通过数据挖掘和分析,建立学习知识模型,为学习者提供详细的学习反馈和建议。

### 2. 学习者行为建模

研究者通过采集学习者在网络学习系统中花费的学习时间、学习者完成课程学习情况、学习者在课堂或学校情境中学习行为变化情况、学习者线上或线下考试成绩等数据,探索学习者学习行为与学习者学习结果的相关关系,最终构建学习者学习行为模型。例如,麦克费登(Macfoyden)和道森(Dawson)通过跟踪和采集 Blackboard 在线学习平台上学习者学习行为相关数据,构建学习者行为模型,该模型被用于预测平台中学习者的学习失败可能。经过实证研究,该模型的预测正确率可以达到 80%以上。

### 3. 学习者经历建模

研究者通过采集学习者的学习满意度调查问卷或量表数据,以及其在后续单元或课程学习中的选择、行为、表现和留存率数据,构建学习者体验模型,利用该模型对在线学习系统中的课程和功能进行评估。例如,可汗学院(Kaplan Inc)通过构建学习者体验模型,对其线上课程进行评估,进行线上课程的再设计,改变课程学习顺序,大大提高了教师的教学效率和学习者的学习成绩。

### 4. 学习者建档

研究者采集在线学习系统中学习者相关数据以及线下

学习者基本信息数据,通过数据挖掘和机器学习算法,构建学习者个人学习档案,分析学习者的学习特征,对具有相同学习特征的学习者进行聚类 and 分组,最终为不同类型的学习者提供个性化的学习环境,促进学习者有效学习的发生。例如,卡丹(Kardan)和克纳蒂(Conati)利用数据挖掘技术对学习者在在线学习平台的交互日志进行分析,确定不同学习者的学习类型和交互特点,构建学习者的学习档案。该研究结果主要用于对新学习者进行分类,并提供合适的学习支持和交互支持。

### 5. 领域知识建模

研究者通过对教育大数据的挖掘和学习分析,对现有领域知识模型进行重构,探索课程、学习单元和知识点的学习内容组织方式与学习者学习结果之间的相关关系。例如,马丁(Martin)等人采集、处理学习者相关数据,构建学习者的学习曲线,并通过对大量学习者学习曲线数据的分析,对现有领域知识模型进行重构。

### 6. 学习组件分析和教学策略分析

研究者通过对学习者在在线学习系统中的学习相关数据进行采集和分析,探索在线学习系统中学习组件的功能、在线教学策略与学习者学习结果的相关关系,进而实现对在线学习系统的评估。例如,里特(Ritter)等人通过对一个名为“Cognitive Tour”的数学教学智能导学系统中数据的挖掘和分析,为该系统提供了长达 15 年的动态、细粒度的系统评价,优化了系统的组成模块和线上教学策略。

### 7. 趋势分析

研究中通过对大量学习者(2000人以上)一段时间内学习相关数据的采集和分析,探索学习者在这一阶段学习过程中的学习结果改变,发现学习者当前学习行为和未来学习结果的相关关系。并利用已建立的相关关系,依据新学习者的当前学习行为,预测其未来的学习趋势和结果。例如,加州高等教育协会为州内高等教育机构提供了一个在线趋势分析工具,该工具允许用户检索加州高等教育数据库,自定义检索条件,生成个性化的教育趋势预测结果。

### 8. 自适应学习系统和个性化学习

这部分应用属于教育数据挖掘和学习分析的高级应用,是教育大数据相关研究的终极目标,通过对教育大数据的采集、处理和分析,最终实现学习者自适应和个性化学习环境的构建。

### (三)教育大数据应用挑战

大数据时代已经来临,大数据必将改变教育的传统面貌,大数据相关理论和技术在教育领域有着广阔的应用前景,如前所述的诸多应用领域,目前由于种种限制,部分大数据的教育应用还处于研究和探索阶段,在《报告》的最后部分,即对这些应用限制和挑战作了详细的说明。

### 1. 技术挑战

大数据教育应用的技术挑战主要有以下三个方面:第一,众所周知,大数据的应用基础就是对海量数据的拥有,这就涉及数据存储技术的挑战,以及用于数据处理和分析的技术挑战,这包括计算机硬件的数据处理能力、超级计算机算法技术等。《报告》通过对技术专家的访谈了解到,他们对近期克服如上所述的技术挑战不持乐观态度。第二,大数据的教育应用中,数据采集和问题解决分析是核心环节,应用开发者要面对的就是数据采集技术和问题解决分析技术的挑战。《报告》中受访专家表示:“如果有 100 人投入到教育大数据的应用工作中,那么需要其中的 99 人投入到数据采集技术和问题解决分析技术的研究工作中<sup>[14]</sup>。”第三,数据兼容性挑战,不同数据存储系统中的数据编码和格式的不统一,造成不同系统间的数据共享困难,导致这一问题的主要原因是各个系统建设和购买缺乏统一规划,无法形成统一数据平台。

### 2. 组织体制挑战

《报告》中受访专家指出,大数据教育应用中所面临的技术挑战,可以通过技术研发和基础设施建设去克服,而大数据教育应用过程有 80% 以上的挑战和限制来源于“人”。由于联邦政府、州和学校之间组织体制的不同,以及不同教师和学校对于教育大数据应用期望的差别,导致应用开发和研究开发人员工作困难,造成许多大数据的教育应用难以实现。《报告》指出,智慧的数据使用者,将会使大数据的应用开发事半功倍,他们可以帮助开发者准确地定位需要解决的问题,确定需要采集的数据类型以及确定有效的可视化数据呈现方式。另外,他们还可以帮助解释数据,判别和标记模式,以及指导模型构建。

### 3. 隐私和伦理道德限制

隐私和伦理道德限制也是大数据教育应用过程中不可避免的问题。《报告》指出,教育机构在将学生的个人教育记录数据交付给第三方机构用于教育数据挖掘和学习分析应用开发的时候,必须注意规避相应的政策和法律规定,美国的《家庭教育权利和隐私法案》是一项联邦法案,其中明确规定要保护学生个人学习记录的隐私。另外,通过教育数据挖掘和学习分析所作出的预测和推荐,将会大大增加学生活动的透明性,这也会带来一系列的社会伦理道德问题,值得相关研究和教育机构重视。

## 四、反思与启示

《报告》作为美国教育部发布的美国国内大数据教育应用的纲领性报告,必将引领美国 K-12 学校和高等教育机构对于教育大数据的研究和应用热潮。基于数据的教学行为优化和学习行为监测并不是新鲜事物,但教育数据挖掘和学习

分析技术的出现,可以实现对于大量细粒度教育大数据的采集、处理和分析,必将改变传统教育数据应用的范式,也将为教育机构的教育决策、教师的教学决策和学生的自我学习监控提供更加精细化的服务。

通过对这份重要报告的解读,我们对于美国近年来大数据的教育应用情况有了一个清晰的了解,既让我们看到了美国在大数据教育应用方面取得的成绩,也帮助我们了解到了美国大数据教育应用目前面临的挑战和困难,对我国未来大数据教育应用的研究和实践有如下启示:

### (一) 宏观层面

宏观层面上,《报告》对我们的启示主要包括国家或地方对于大数据教育应用研究和实践的技术支撑、管理体制支撑和法律制度支撑。

#### 1. 技术支撑

(1) 我国目前大数据相关的技术资源与美国相比还略显薄弱,区域间技术资源分布不均,很多地方教育机构缺乏必要的大数据应用的基础设施建设。例如,某些学校没有学生信息系统、在线教育平台,也就无法为教育数据挖掘和学习分析提供基础数据,这将是我国在大数据教育应用方面所面临的首要问题。面对此类问题,结合我国的具体国情,我们要发挥后发优势,利用云计算和虚拟化等先进信息技术,实现大数据应用相关的技术资源的跨越式部署。

(2) 数据永远是大数据应用的核心,数据兼容性问题是在推进大数据教育应用过程中所要重点关注的问题。我国目前已有的学生信息系统、在线教育平台等大数据来源系统也存在数据结构和数据格式不统一的问题,各级教育机构在这些系统的选择方面缺乏统一规划和全局考虑。因此,在未来的建设中,我们对于类似系统的规划和建设要采取顶层设计的原则,由教育部对类似系统的数据格式和数据结构指定统一的规范,各级各类教育机构在进行类似系统的规划和建设中,系统的数据结构和数据格式要符合统一规范要求,使得各级教育机构的系统数据可以实现无障碍迁移和共享。

(3) 在大数据教育应用初期,研究机构和教育机构往往过分关注技术本身,而忽视教育实际,造成本末倒置,舍本逐末的情况出现。因此,在规划教育大数据挖掘和学习分析具体应用的时候,一定要按照解决问题界定到分析数据选择,再到技术选择的技术路线进行,真正做到有的放矢,避免不必要的人力和物力的浪费。

#### 2. 管理体制支撑

(1) 在国家方面,不仅要制定大数据教育应用的远景规划,还要尽快出台实施细则,以指导各级各类教育机构推进大数据教育应用。另外,必要的情况下可以成立相应的组织机构,其职能主要负责统一数据规范制定,技术方案推荐以及组织协调。



(2)在各级教育机构方面,学生信息系统、在线教育平台等大数据来源系统的规划和采购,要充分考虑大数据的应用实际,确保系统的数据结构和格式符合统一规范,能够为基于大数据分析的教学改善和教育决策制定提供充足的数据支持。另外,在具体的教育大数据应用规划时,要摒弃以技术部门决策为主的思想,确立由技术部门、管理部门和教学部门共同参与的联合决策为主。在具体的应用实施策略方面,可以按照从小做起,边总结边推进的思路进行,例如,一个学校可以先从本校的一个系统的数据挖掘和分析做起,然后尝试整合校内其他系统的数据挖掘和分析,接下来联合区域内学校开展区域教育大数据的应用,最终实现与国家层面的教育大数据应用的对接。

(3)在教育管理者和教师方面,无论是地区教育管理部门的管理者,还是学校管理者,在制定任何教育决策的时候,都要牢固树立基于数据分析的科学决策理念,教师在教学内容和教学策略的选择方面,也要基于数据分析进行。

### 3.法律制度支撑

这一方面在大数据教育应用的初期往往容易被人们忽视,我们在看到大数据给我们的教育带来改变的同时,也应该注意到在大数据教育应用实施过程中所带来的诸多法律制度和伦理道德问题。因此,在大数据教育应用的规划时就应该充分考虑到所需采集数据涉及的个人隐私问题,以及由大数据应用造成学生活动透明化所带来的伦理道德问题,制定和完善相应的法律法规及具体制度,避免由于技术应用而造成深层次的社会问题。

## (二)微观层面

在微观层面上,《报告》对我们的启示主要是未来在教学和学习方面,基于大数据可以规划的具体的研究和应用:

第一,大数据时代的虚拟学习社区构建。大数据时代的虚拟学习社区呈现出关联性更强的社会网络结构,这就使得我们在虚拟学习社区的构建和运维等方面面临机遇和挑战。如何通过大数据挖掘和分析,探索社会网络位置、资源分配和知识建构等变量的相关关系,建立相应的分析和预测机制,为虚拟学习社区的构建和运维提供科学有效的决策参考,值得我们未来展开研究。另外,大数据时代的虚拟学习社区,随着数字学习资源的不断积累,不可避免地要面临数字资源的保护问题,相应的保护机制、技术和方法也将是未来研究的主要方向。目前,国内已有学者针对这一问题展开了相关研究,例如,李凤英等人提出了一种基于门限代理签名的虚拟学习社区数字资源保护方案<sup>[5]</sup>。

第二,基于大数据挖掘和分析的学习资源推送机制研究。大数据时代数字化的学习资源将呈现爆炸式的增长,面对海量的学习资源,学习者如何选择,或者如何向学习者推送合适的学习资源将成为未来面临的主要问题。基于大数据

的挖掘和分析,可以精细分析学习者的学习风格、特点和现状,为学习者主动推送合适的学习资源,相关的技术、方法和机制值得我们在未来展开研究。

第三,基于大数据挖掘和分析的学习者学习路径优化研究。大数据时代为我们提供的海量数据和先进的数据分析技术,使得我们有机会更加精细地去认识每一位学习者,通过对学习者相关数据的挖掘和分析,绘制学习者学习路径图,在前期预测模型研究的基础上,优化学习者的学习路径,促进学习反思,与之相关的技术和方法值得我们在未来展开研究。

## 五、结语

综上所述,我们已经进入了一个“数据驱动学校,分析变革教育”的大数据时代,大数据必将改变传统教育的面貌,“数据驱动决策”已经成为教育流行语。发达国家已经将大数据提升到国家战略层面,并在大数据相关领域投入巨资,教育领域的大数据相关研究和应用已经取得了一定的成果。而我们国家尚处启蒙阶段,因此,我们迫切需要整合现有资源,发挥后进优势,实现教育大数据研究和应用的跨越式发展。借助大数据,我们将更好地读懂千差万别的学生,让我们有机会了解每一个学生的真实学习情况,为每一个学生提供优质、个性化的教育资源,不断实现教育公平。

## [参考文献]

- [1]Big Data for Development: Challenges & Opportunities [DB/OL]. [2012-05-01]. <http://www.unglobalpulse.org/sites/default/files/Big-DataforDevelopment-UNGlobalPulseJune2012.pdf>.
- [2]Big Data Research and Development Initiative[DB/OL].[2012-03-29]. [http://www.whitehouse.gov/sites/default/files/microsites/ostp/big\\_data\\_press\\_release\\_final\\_2.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf).
- [3][7][12][13][14] Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics [DB/OL]. [2012-10-12]. <http://www.ed.gov/edblogs/technology/files/2012/03/edm-la-brief.pdf>.
- [4]Big data-Wikipedia, the free encyclopedia [EB/OL].[2013-09-23]. [http://en.wikipedia.org/wiki/Big\\_data](http://en.wikipedia.org/wiki/Big_data).
- [5]Barwick H. The "four Vs" of Big Data. Implementing Information Infrastructure Symposium[EB/OL]. [2012-10-02]. [http://www.computer-world.com.au/article/396198/iiis\\_four\\_vs\\_big\\_data/](http://www.computer-world.com.au/article/396198/iiis_four_vs_big_data/).
- [6]Manyika. J. Big data: The next frontier for innovation, competition, and productivity [EB/OL]. [2011-05-12]. [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation).
- [8]Siemens, G. Learning and Knowledge Analytics -Knewton -the future of education?[EB/OL]. [2011-04-14]. <http://www.learninganalytics.net/?p=126>.



[9]NMC Horizon Repor 2011 Higher Ed Edition [EB/OL].[2011-01-10].  
http://www.nmc.org/publications/horizon-report-2011-higher-ed-edition.

[10]顾小清,张进良,蔡慧英.学习分析:正在浮现中的数据技术[J].远程教育杂志,2012,(1):18-25.

[11]李艳燕,马韶茜,黄荣怀.学习分析技术:服务学习过程设计和优化[J].开放教育研究,2012,(5):18-24.

[15]李凤英,齐宇歆,薛庆水.大数据视域下的虚拟学习社区安全研究

[J].远程教育杂志,2013,(4):76-82.

#### [作者简介]

徐鹏,东北师范大学在读博士研究生,工程师,研究方向:教师教育技术、大数据教育应用;王以宁,教授,博士生导师,研究方向:教师教育技术;刘艳华,吉林大学公共外语教育学院讲师,研究方向:网络教育、数字化教学设计;张海,东北师范大学传媒科学学院副教授,博士生导师,研究方向:教师教育技术。

### The Learning Innovation from the Perspective of Big Data: An Analysis of the U.S. Report of *Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics* and Its Enlightenment

Xu Peng<sup>1</sup>, Wang Yining<sup>2</sup>, Liu Yanhua<sup>3</sup> & Zhang Hai<sup>2</sup>

(1. Informatization Office, Northeast Normal University;

2. School of Media Science, Northeast Normal University;

3. School of Foreign Language Education, Jilin University, Jilin, Changchun 130012)

【Abstract】Ready or not, here comes the big data era when technologies on educational data mining and learning analytics are adopted, relevant models in the educational field are constructed, the relationships between educational variables are explored, and the decision-making in education and teaching is strongly supported. Big Data has been witnessing another innovation of human learning. In 2012, the Department of Education of the U.S. issued a report named *Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics*, in which the application fields, cases and challenges of the implementation of the Big Data education were thoroughly presented. In the future, the Big Data research and application in the educational field in China should focus on the national and local support from the technical level, the management level and the legislative level, through which specific research and application should be planned rationally on demand, existing resources integrated, the late-development advantages given full play to, and thus, with the help of Big Data, educational equity can be achieved through the personalized learning in real sense.

【Keywords】Big data; Educational data mining; Learning analytics

收稿日期:2013年10月23日

责任编辑:刘菊

## 浙江省首届高校微课比赛颁奖仪式 暨微时代高校教学信息化研讨会在杭州师大召开

【本刊讯】2013年10月26日,由教育部全国高校教师网络培训中心杭州分中心主办、杭州师范大学承办的“浙江省首届高校微课大赛颁奖暨微时代教学信息化研讨会”在杭州师范大学仓前校区召开。

会上,来自华南师大焦建利教授作了“大规模在线开放课程”的专题报告。教育部网培中心的段露航博士以“微课与高校教学改革”为题,从网培中心的角度对微课大赛的背景以及在促进教学改革方面可能发挥的作用进行了阐述。杭州师范大学外国语学院杨小洪教授

做了“混合式课堂教学实践”的主题报告。此外,大会还分享了上海高校课程共享中心教学模式、高校教育信息化的建设经验。

会议分本科组和高职组两个分会场,分别进行作品的展示、点评与经验分享。此次研讨会为浙江省高校教师展示教学技能搭建了一个很好的交流平台,进一步促进了浙江省高校优质教育资源的共建共享,推动全省高校教育技术工作的不断发展。

(杭技)